



Improving mesoscopic fluorescence molecular tomography via preconditioning and regularization

FUGANG YANG,^{1,*} RUOYANG YAO,² MEHMET OZTURK,² DENZEL FAULKNER,² QINGLAN QU,³ AND XAVIER INTES²

¹School of Information and Electronic Engineering, Shandong Institute of Business and Technology, Yantai 264005, China

²Department of Biomedical Engineering, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180, USA

³Department of Reproductive Medicine, Yantai Yuhuangding Hospital, Affiliated Hospital of Qingdao University, Yantai 264000, China

*Fugang@sdibt.edu.cn

Abstract: Mesoscopic fluorescence molecular tomography (MFMT) is a novel imaging technique capable of obtaining 3-D distribution of molecular probes inside biological tissues at depths of a few millimeters with a resolution up to $\sim 100\ \mu\text{m}$. However, the ill-conditioned nature of the MFMT inverse problem severely deteriorates its reconstruction performances. Furthermore, dense spatial sampling and fine discretization of the imaging volume required for high resolution reconstructions make the sensitivity matrix (Jacobian) highly correlated, which prevents even advanced algorithms from achieving optimal solutions. In this work, we propose two computational methods to respectively increase the incoherence of the sensitivity matrix and improve the convergence rate of the inverse solver. We first apply a compressed sensing (CS) based preconditioner on either the whole sensitivity matrix or sub sensitivity matrices to reduce the coherence between columns of the sensitivity matrix. Then we employed a regularization method based on the weight iterative improvement method (WIIM) to mitigate the ill-condition of the sensitivity matrix and to drive the iterative optimization process towards convergence at a faster rate. We performed numerical simulations and phantom experiments to validate the effectiveness of the proposed strategies. In both *in silico* and *in vitro* cases, we were able to improve the quality of MFMT reconstructions significantly.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

OCIS codes: (170.3010) Image reconstruction techniques; (170.3880) Medical and biological imaging; (100.3190) Inverse problems; (170.6960) Tomography; (170.2520) Fluorescence microscopy.

References and links

1. E. M. C. Hillman, D. A. Boas, A. M. Dale, and A. K. Dunn, "Laminar optical tomography: demonstration of millimeter-scale depth-resolved imaging in turbid media," *Opt. Lett.* **29**(14), 1650–1652 (2004).
2. M. S. Ozturk, V. K. Lee, L. Zhao, G. Dai, and X. Intes, "Mesoscopic fluorescence molecular tomography of reporter genes in bioprinted thick tissue," *J. Biomed. Opt.* **18**(10), 100501 (2013).
3. M. S. Ozturk, C. W. Chen, R. Ji, L. Zhao, B. N. Nguyen, J. P. Fisher, Y. Chen, and X. Intes, "Mesoscopic Fluorescence Molecular Tomography for Evaluating Engineered Tissues," *Ann. Biomed. Eng.* **44**(3), 667–679 (2016).
4. M. S. Ozturk, X. Intes, G. dai, and V. K. Lee, "Longitudinal Volumetric Assessment of Glioblastoma Brain Tumor in 3D Bio-Printed Environment by Mesoscopic Fluorescence Molecular Tomography," in *Biomedical Optics 2016*, OSA Technical Digest (online) (Optical Society of America, 2016), paper JM3A.46.
5. Q. Tang, J. Lin, V. Tsytarev, R. S. Erzurumlu, Y. Liu, and Y. Chen, "Review of mesoscopic optical tomography for depth-resolved imaging of hemodynamic changes and neural activities," *Neurophotonics* **4**(1), 011009 (2016).
6. M. Azimipour, M. Sheikhzadeh, R. Baumgartner, P. K. Cullen, F. J. Helmstetter, W. J. Chang, and R. Pashaie, "Fluorescence laminar optical tomography for brain imaging: system implementation and performance evaluation," *J. Biomed. Opt.* **22**(1), 16003 (2017).

7. F. Long and X. Intes, "Dental optical tomography with upconversion nanoparticles—a feasibility study," *J. Biomed. Opt.* **22**(6), 66001 (2017).
8. M. S. Ozturk, D. Rohrbach, U. Sunar, and X. Intes, "Mesoscopic fluorescence tomography of a photosensitizer (HPPH) 3D biodistribution in skin cancer," *Acad. Radiol.* **21**(2), 271–280 (2014).
9. Y. Chen, S. Yuan, J. Wierwille, R. Naphas, Q. Li, T. R. Blackwell, T. W. Paul, Jr., and G. Kristine, "Integrated optical coherence tomography (OCT) and fluorescence laminar optical tomography (FLOT)," *IEEE J. Sel. Top. Quant.* **16**(4), 755–766 (2010).
10. D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proc. Natl. Acad. Sci. U.S.A.* **100**(5), 2197–2202 (2003).
11. D. L. Donoho, Y. Tsaig, I. Drori, and J. L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory* **58**(2), 1094–1121 (2012).
12. M. T. Schaub, M. Trefois, P. van Dooren, and J.-C. Delvenne, "Sparse matrix factorizations for fast linear solvers with application to Laplacian systems," *SIAM J. Matrix Anal. Appl.* **38**(2), 505–529 (2017).
13. M. Benzi, "Preconditioning techniques for large linear systems: a survey," *J. Comput. Phys.* **182**(2), 418–477 (2002).
14. F. Yang, M. S. Ozturk, L. Zhao, W. Cong, G. Wang, and X. Intes, "High-resolution mesoscopic fluorescence molecular tomography based on compressive sensing," *IEEE Trans. Biomed. Eng.* **62**(1), 248–255 (2015).
15. Q. Pian, R. Yao, N. Sinsuebphon, and X. Intes, "Compressive hyperspectral time-resolved wide-field fluorescence lifetime imaging," *Nat. Photonics* **11**(7), 411–414 (2017).
16. A. Jin, B. Yazici, and V. Ntziachristos, "Light illumination and detection patterns for fluorescence diffuse optical tomography based on compressive sensing," *IEEE Trans. Image Process.* **23**(6), 2609–2624 (2014).
17. A. Jin, B. Yazici, A. Ale, and V. Ntziachristos, "Preconditioning of the fluorescence diffuse optical tomography sensing matrix based on compressive sensing," *Opt. Lett.* **37**(20), 4326–4328 (2012).
18. R. Yao, Q. Pian, and X. Intes, "Wide-field fluorescence molecular tomography with compressive sensing based preconditioning," *Biomed. Opt. Express* **6**(12), 4887–4898 (2015).
19. F. Yang, M. S. Ozturk, R. Yao, and X. Intes, "Improving mesoscopic fluorescence molecular tomography through data reduction," *Biomed. Opt. Express* **8**(8), 3868–3881 (2017).
20. V. L. Wang, and I. W. Hsin *Biomedical optics: principles and imaging* (John Wiley & Sons, 2012).
21. C. Zhu and Q. Liu, "Review of Monte Carlo modeling of light transport in tissues," *J. Biomed. Opt.* **18**(5), 50902 (2013).
22. L. Wang and S. L. Jacques, "Hybrid model of Monte Carlo simulation and diffusion theory for light reflectance by turbid media," *J. Opt. Soc. Am. A* **10**(8), 1746–1752 (1993).
23. S. H. Tseng, C. Hayakawa, J. Spanier, and A. J. Durkin, "Investigation of a probe design for facilitating the uses of the standard photon diffusion equation at short source-detector separations: Monte Carlo simulations," *J. Biomed. Opt.* **14**(5), 054043 (2009).
24. Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Opt. Express* **17**(22), 20178–20190 (2009).
25. J. Chen and X. Intes, "Comparison of Monte Carlo methods for fluorescence molecular tomography—computational efficiency," *Med. Phys.* **38**(10), 5788–5798 (2011).
26. J. Chen, V. Venugopal, and X. Intes, "Monte Carlo based method for fluorescence tomographic imaging with lifetime multiplexing using time gates," *Biomed. Opt. Express* **2**(4), 871–886 (2011).
27. S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1 -regularized least squares," *IEEE J. Sel. Top. Signal Process.* **1**(4), 606–617 (2007).
28. J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.* **18**(7), 1395–1408 (2009).
29. Y. Saad and H. A. van der Vorst, "Iterative solution of linear systems in the 20th century," *J. Comput. Appl. Math.* **123**(1–2), 1–33 (2000).
30. X. Deng, L. Yin, S. Peng, and M. Ding, "An iterative algorithm for solving ill-conditioned linear least squares problems," *Geod. Geodyn.* **6**(6), 453–459 (2015).
31. J. A. Tropp, "Greed is Good: Algorithmic Results for Sparse Approximation," *IEEE Trans. Inf. Theory* **50**(10), 2231–2242 (2004).
32. L. Zhao, V. K. Lee, S.-S. Yoo, G. Dai, and X. Intes, "The integration of 3-D cell printing and mesoscopic fluorescence molecular tomography of vascular constructs within thick hydrogel scaffolds," *Biomaterials* **33**(21), 5325–5332 (2012).
33. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**(11), R37–R61 (2013).
34. L. Zhao, H. Yang, W. Cong, G. Wang, and X. Intes, "Lp regularization for early gate fluorescence molecular tomography," *Opt. Lett.* **39**(14), 4156–4159 (2014).
35. V. Venugopal, J. Chen, M. Barroso, and X. Intes, "Quantitative tomographic imaging of intermolecular FRET in small animals," *Biomed. Opt. Express* **3**(12), 3161–3175 (2012).
36. K. Abe, L. Zhao, A. Periasamy, X. Intes, and M. Barroso, "Non-invasive in vivo imaging of near infrared-labeled transferrin in breast cancer cells and tumors using fluorescence lifetime FRET," *PLoS One* **8**(11), e80269 (2013).
37. S. Rajoria, L. Zhao, X. Intes, and M. Barroso, "FLIM-FRET for cancer application," *Curr. Mol. Imaging* **3**(2), 144–161 (2015).

38. X. Intes, B. Chance, M. Holboke, and A. Yodh, "Interfering diffusive photon-density waves with an absorbing-fluorescent inhomogeneity," *Opt. Express* **8**(3), 223–231 (2001).
39. Y. Chen, C. Mu, X. Intes, D. Blessington, and B. Chance, "Near-infrared phase cancellation instrument for fast and accurate localization of fluorescent heterogeneity," *Rev. Sci. Instrum.* **74**(7), 3466–3473 (2003).
40. X. Intes, V. Ntziachristos, and B. Chance, "Analytical model for dual-interfering sources Diffuse Optical Tomography," *Opt. Express* **10**(1), 2–14 (2002).
41. V. A. Markel, V. Mital, and J. C. Schotland, "Inverse problem in optical diffusion tomography. III. Inversion formulas and singular-value decomposition," *J. Opt. Soc. Am. A* **20**(5), 890–902 (2003).

1. Introduction

As a novel imaging modality combining the benefits of millimeter deep molecular imaging with high sensitivity, Mesoscopic Fluorescence Molecular Tomography (MFMT) can resolve the distribution and concentration of fluorophores inside biological tissues based on boundary measurements [1,2]. MFMT has the unique potential to cover the gap between microscopic and macroscopic examination of biotissues by enabling resolution up to hundred micrometers level at a depth of a few millimeters in intact tissues (without optical windows or clearing agents). MFMT has already found applications in imaging bio-printed thick constructs [3,4], brain molecular [5,6], dental imaging [7] and tumor xenografts studies [8,9]. However, performances of MFMT are intrinsically associated with the inverse optical problem that need to be solved.

Typically, the MFMT inverse problem is ill-posed and ill-conditioned. Hence it is very challenging to solve it in a robust fashion. Among the various techniques developed to solve ill-conditioned problems, iterative methods, embedded with various optimization algorithms, have become an effective approach that use successive approximations to obtain solutions with the help of additional constraints introduced by appropriate regularization techniques [10,11]. However, the performance of the iterative method adopted depends greatly on the incoherence and the spectrum of the forward sensitivity matrix. Additionally, many iterative methodologies rely on preconditioners to improve performance and ensure fast convergence [12]. In the case of the MFMT inverse problem, the measurements acquired are characterized with high redundancy due to the diffuse nature of the light collected, reflection geometry employed and dense spatial sampling. Hence, the associated Jacobian demonstrates high coherence between its columns and high condition number due to fine discretization of the volume to be imaged to attain hundred micrometers resolution. In turn, the designed iterative optimization algorithm will slowly converge and may be susceptible to the propagation and amplification of errors after successive iterations [13].

On the other hand, Compressed Sensing (CS) methodologies have been proved to be successful methods to recover sparse signals from far fewer samples than required by the traditional Shannon-Nyquist sampling theorem [14,15]. The CS framework guarantees accurate recovery of sparse signals under certain conditions, one of which is incoherence, *i.e.* in our case, the degree of the orthogonality of forward sensitivity matrix. Recent works in diffuse optical tomography (DOT) have demonstrated that using a properly preconditioning matrix can reduce the coherence between columns of the forward sensitivity matrix, and as a consequence alleviate the ill-posedness of the underdetermined linear system [16,17]. Similarly, our group has proposed a preconditioning strategy to reduce the coherence of sensitivity matrix for wide-field fluorescence molecular tomography [18]. As for the MFMT inverse problem itself, Yang et al. [19] investigated the influence of noise on reconstruction performance and addressed a two-step data reduction approach to achieve high fidelity results. Nevertheless, the implementation of effective preconditioning methodologies in MFMT has not been reported to date but is expected to greatly improve the performances of the reconstruction process.

In this paper, we propose two computational methods to improve both incoherence and condition number of the sensitivity matrix of MFMT. First, we propose a preconditioner to reduce the correlation between columns of the Jacobian, which contribute to mitigating the ill-posedness of the inverse problem. Then another well-designed regularization parameter is

added to improve the convergence rate of the inverse solver to the targeted tolerance. The proposed strategies are described in Section 2 in details. Section 3 presents the objective metrics to evaluate spectrum property, coherence of Jacobian, noise level of measurements, and the MFMT reconstruction performances. Section 4 summarizes the *in silico* and *in vitro* experiment setups with reconstruction results. The discussion and conclusion are provided in Section 5.

2. Method

2.1 Configuration of the MFMT optical system

The optical configuration of our second generation MFMT system has been described in details in previous reports [4,19]. Here, we provide the main salient features that are relevant to the optical inverse problem. Briefly, the system is built around the raster scanning of an illumination spot over the specimen with detection performed by an electron-multiplying CCD camera (emCCD) (iXon^{EM+} DU-897 back illuminated, Andor Technology) acquiring data in a descanned configuration. This set up enables the acquisition of up to 512×512 measurements in parallel per illumination spot, leading to very dense spatial data sets. To obtain a data set that can be inverted efficiently, the emCCD data can be binned into super-pixels that still provide high accuracy and resolution [19]. The discretization of the image space and positioning of the optodes configuration are illustrated in Fig. 1, where the green squares, red squares and blue squares respectively represent the discretized voxels of the imaging volume, pixels of emCCD camera, and super-pixels binned at 2×2 to provide enhanced SNR.

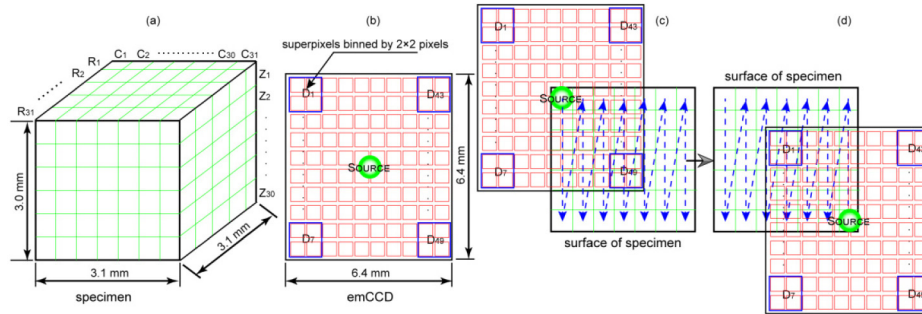


Fig. 1. Schematic diagram of scanning mode of the 2nd generation MFMT system. (a) displays the discretization of imaging volume. (b) shows the dimension of the emCCD with super pixels (blue squares) binned by 2×2 pixels (red squares) used as detectors. (c) and (d) illustrate the scanning trajectory (blue dash) from start scanning spot (c) to the last one (d).

This configuration and on the chip binning strategy leads to a total of 256×256 super pixels as detectors that cover a detection FOV of $6.4 \times 6.4 \text{ mm}^2$. To cast the inverse problem, we further down sample the measurements space by selecting a 7×7 detector matrix (D_1, D_2, \dots, D_{49}) with a separation of $\sim 0.6 \text{ mm}$. To improve the dynamical range, the central detection which coincides with the position of the illumination spot is occluded during acquisition so detector 25 (D_{25}) is not employed. On the illumination side, the raster scanning step size and dwelling time are typically set to $100 \mu\text{m}$, 20 ms respectively, leading to a total of 961 scanning locations in the illumination FOV of $3.1 \times 3.1 \text{ mm}^2$. Last, the volume to be imaged is discretized uniformly in voxels of $100 \mu\text{m}$ length, leading to R_i ($i = 1, \dots, 31$), C_j ($j = 1, \dots, 31$), and Z_k ($k = 1, 2, \dots, 30$) number of voxels along the x, y and z axes respectively. Overall, this configuration leads to 46,128 measurements that are employed in the inverse optical problem (out of 251,920,384 possible without binning).

2.2 Forward model and inverse problem

The source-detector separation in our MFMT implementation can be as small as 600 μm . Hence, in view of typical optical properties of tissues, the collected photons cannot be accurately modeled by the Diffusion Equation [20]. Therefore, we employ Monte Carlo (MC) method, commonly known as the gold standard [21]. The MC method is well established as being accurate in modeling photon propagation in shallow tissues and near sources [22,23]. To meet the accuracy requirements under high resolution of MFMT operating in the mesoscopic regime, up to 10^8 photons are simulated for each source/detector with the GPU-based software MCX to simulate continuous-wave (CW) Green's functions G^x and G^m at excitation and emission wavelengths [24]. The weight function W for one source-detector pair is then computed with an adjoint formulation for efficiency [25]:

$$W(r_s, r_d, r) = G^x(r_s, r) \times G^m(r, r_d) \quad (1)$$

where r_s , r_d , and r are locations for source, detector and any location inside the specimen being probed, respectively. The measured fluorescence intensity for the corresponding source-detector pair $U(r_s, r_d)$ can be formulated as the integral equation below:

$$U(r_s, r_d) = \int_{\Omega} W(r_s, r_d, r) \eta(r) dr^3 \quad (2)$$

where $\eta(r)$ is the 3-D distribution of the fluorophore's effective quantum yield. As at this time our experimental system is limited to the continuous wave (CW) data type for MFMT applications, however it's straightforward to expand both Eqs. (1,2) to time-dependent form given time-resolved Green's functions and the lifetime of fluorophore τ [26]. With M source-detector pairs, and N discretized voxels of the Region of Interest (ROI), the inverse problem for MFMT in the matrix form becomes:

$$AX = b \quad (3)$$

where $A \in \mathbb{R}^{M \times N}$ is the Jacobian, $X \in \mathbb{R}^{N \times 1}$ represents the effective quantum yield, $b \in \mathbb{R}^{M \times 1}$ is the vector of detector readings at all source positions.

The CS based framework has been demonstrated as an effective methodology to exactly retrieve sparse solutions from under-determined linear systems [10,11,14]. Fortunately, the solutions in our MFMT application are inherently sparse because fluorescent biomarkers are supposed to be clustered within small regions of tissues such as tumors. To achieve sparsity in the image space, we add the commonly used l_1 -norm regularization term so that the optimization problem now becomes:

$$\min \{ \|AX - b\|_2^2 + \lambda \|X\|_1 \} \quad (4)$$

where $\|X\|_1 = \sum_{i=1}^n |x_i|$ denotes the l_1 -norm of fluorophore distribution vector and λ is the regularization parameter. The minimization problem in Eq. (4) can then be transformed to a convex quadratic problem with linear inequality constraints, and solved by a variety of inverse solvers such as the interior-point method that we applied in this work [27].

2.3 Preconditioning to reduce the coherence of the sensitivity matrix

The first preconditioning strategy applied aims at reducing the coherence of the sensitivity matrix, so that the performance of CS-based sparse signal recovery algorithms is improved. The design of preconditioner follows a similar approach as described in [16–18] so we only briefly introduce the derivation. When a preconditioner M_A is applied, the original Jacobian matrix becomes:

$$A_{\text{pre}} = M_A A \quad (5)$$

For simplicity, we assume that the preconditioned sensitivity matrix \mathbf{A}_{pre} is also column-normalized, and to minimize the coherence of \mathbf{A}_{pre} , we seek to determine \mathbf{M}_A such that the Gramian matrix $\mathbf{A}_{\text{pre}}^T \mathbf{A}_{\text{pre}}$ approximates the identity matrix [10,11,28],

$$\mathbf{A}_{\text{pre}}^T \mathbf{A}_{\text{pre}} = \mathbf{A}^T \mathbf{M}_A^T \mathbf{M}_A \mathbf{A} \approx \mathbf{I} \quad (6)$$

Equation (6) still holds after multiplying \mathbf{A} on the left and \mathbf{A}^T on the right,

$$\mathbf{A} \mathbf{A}_{\text{pre}}^T \mathbf{A}_{\text{pre}} \mathbf{A}^T = \mathbf{A} \mathbf{A}^T \mathbf{M}_A^T \mathbf{M}_A \mathbf{A} \mathbf{A}^T \approx \mathbf{A} \mathbf{A}^T \quad (7)$$

Taking the Singular Value Decomposition of $\mathbf{A} = \mathbf{U}_A \mathbf{\Sigma}_A \mathbf{V}_A^T$, and substitute it into Eq. (7), we get:

$$(\mathbf{\Sigma}_A \mathbf{\Sigma}_A^T) \mathbf{U}_A^T \mathbf{M}_A^T \mathbf{M}_A \mathbf{U}_A (\mathbf{\Sigma}_A \mathbf{\Sigma}_A^T) \approx \mathbf{\Sigma}_A \mathbf{\Sigma}_A^T \quad (8)$$

Thus, we choose \mathbf{M}_A as the preconditioner,

$$\mathbf{M}_A = (\mathbf{\Sigma}_A \mathbf{\Sigma}_A^T)^{-\frac{1}{2}} \mathbf{U}_A^T \quad (9)$$

During the actual implementation, a regularization parameter is usually added for stabilization when having a large condition number, so the final form of our first preconditioner becomes:

$$\mathbf{M}_A = (\mathbf{\Sigma}_A \mathbf{\Sigma}_A^T + \epsilon \mathbf{I})^{-\frac{1}{2}} \mathbf{U}_A^T \quad (10)$$

where $\epsilon = 0.1$ is empirically selected to provide the best trade-off between reducing matrix coherence and suppressing noise propagation in the measurements under the experimental noise level (SNR = 4) [18].

Note that although the construction of MFMT Jacobian also follows the adjoint form, we cannot separately reduce the coherence of excitation and emission matrices and compute the preconditioned sensitivity matrix through the Kronecker product, referred as “separate masks” in [18]. This is because the excitation and emission light fields of MFMT are coupled with each other, i.e., the source and CCD camera always move together and remain the same relative position as the scanning proceeds. Thus, only “global mask” strategy can be applied to MFMT, although it’s shown that “separate masks” are slightly superior in suppressing noise and improving resolution in the reconstruction results [16,18]. Though, we can harness the symmetry of the descanned acquisition configuration that leads to a block configuration in the forward model (one source-48 detectors).

Herein, we propose a “sub-preconditioning” approach in which sub-matrices are preconditioned to overall lead to a reduction of the coherence of the whole matrix. In our case, 48 measurements are collected from each of the 961 scanning positions, so we can either perform preconditioning on 961 sub-matrices with 48 rows, or on 48 sub-matrices with 961 rows. In practice, we chose the latter one because it’s more effective at reducing the overall coherence of the Jacobian. Thus, each sub-sensitivity matrix \mathbf{A}_p ($p = 1, \dots, 48$) corresponds to measurements from the same detector at all scanning positions. In the following studies, preconditioning strategies will be performed both globally and block by block, referred as “whole-preconditioning” and “sub-preconditioning”, and the performance of the two approaches are compared and analyzed.

2.4 WIIM to speed-up the convergence rate of the inverse solver

The convergence rate of an iterative optimal algorithm to solve a large linear system depends mainly on the condition number of its coefficient matrix [13,29]. Unfortunately the condition number in MFMT application is usually very large due to the high correlation between

different source-detector pairs resulting from dense sampling and fine discretization. Hence an effective method to reduce the condition number of the sensitivity matrix is highly desired.

To this end, here we employ a regularization approach, Weighted Iterative Improvement Method (WIIM) [30], to effectively improve the ill-conditioned situation and thus contribute to a faster convergence rate of the iterative algorithm. We first construct the iterative process as following,

$$(\Lambda + \gamma \mathbf{I}) \mathbf{x}_{k+1} = \mathbf{y} + \gamma \mathbf{x}_k \quad (11)$$

where $\Lambda = \mathbf{A}^T \mathbf{A}$ is a square matrix, $\mathbf{y} = \mathbf{A}^T \mathbf{b}$, \mathbf{I} is the identity matrix, and $k = 1, 2, 3, \dots$ is the iteration number. The nonzero constant γ determines the degree of improvement for the condition number of Λ . A small γ leads to minor improvement of convergence rate while a very large γ could result in stagnation of iteration or even divergence. It's usually determined from the following empirical formula [30]:

$$\frac{\gamma}{\gamma + \delta^2} \leq 0.1, (\gamma > 0) \quad (12)$$

where δ is the lower bound of all nonzero elements in the sensitivity matrix. The initial value of γ can be given according to Eq. (13)

$$\begin{cases} \gamma = \beta 10^{0.5|\log(\beta)|+1} \\ \beta = \min(|\text{eig}(\Lambda)|) \end{cases} \quad (13)$$

where β is the absolute value of the smallest eigenvalue of matrix Λ and \log is logarithm at the base 10. Let $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{e}_k$ and $\mathbf{r}_k = \mathbf{y} - \Lambda \mathbf{x}_k$, we can substitute them into Eq. (11) and get:

$$(\Lambda + \gamma \mathbf{I}) \mathbf{e}_k = \mathbf{r}_k \quad (14)$$

We can apply the same CS-based solver as we solve Eq. (3) because \mathbf{e}_k is the difference of fluorophore concentration between two iterations and thus is also sparse. With an initial guess \mathbf{x}_0 , we can then repeatedly solve Eq. (14) and update the values of \mathbf{e}_k , \mathbf{r}_k , and \mathbf{x}_k . If the last residual norm $\|\mathbf{r}^k\|$ reaches the targeted tolerance, or the iteration number reaches the maximum, the approximated solution is reached. Note that the iteration steps described above can be easily implemented to be embedded in any selected inverse solver. Last, a positive constraints was applied during the iterative process.

3. Evaluation metrics for an objective assessment of the proposed methods

3.1 The effect of the condition number on the precision of the solution

The condition number of the sensitivity matrix determines how much a small disturbance in the measurement \mathbf{b} could affect the solution of the inverse problem \mathbf{x} in Eq. (3):

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \text{cond}(\mathbf{A}) \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|} \quad (15)$$

where $\text{cond}(\mathbf{A}) = \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$ is the condition number. A smaller condition number could lead to results that are accurate and stable against noise, because the error magnification during iterations is minimized. Therefore, it's beneficial to apply preconditioning methods to reduce the large condition number of the MFMT sensitivity matrix.

3.2 Coherence of the sensitivity matrix

One of the most commonly used metric in the compressed sensing literature to assess the incoherence of a matrix \mathbf{A} is the mutual coherence [16], defined as the largest normalized inner product of two different columns a_p and a_q :

$$M(\mathbf{A}) = \max_{p,q,p \neq q} \frac{|\langle a_p, a_q \rangle|}{\|a_p\|_2 \|a_q\|_2} \quad (16)$$

$M(\mathbf{A})$ will be small only if all the columns of \mathbf{A} are almost orthogonal to each other, so it reports on the worst orthogonality of the whole matrix. Thus, to provide a more robust assessment of the incoherence of the sensitivity matrix, the cumulative coherence or the Babel function can be applied to report on the “average” coherence instead [31].

$$M_1(k, \mathbf{A}) = \max_p \max_{|Q|=k, p \notin Q} \sum_{p \in Q} \frac{|\langle a_p, a_q \rangle|}{\|a_p\|_2 \|a_q\|_2} \quad (17)$$

Where $|Q|$ is the cardinality, i.e. the number of columns, of the selected subset of the sensitivity matrix \mathbf{Q} , and $M_1(k, \mathbf{A})$ is monotonically non-decreasing [31].

3.3 Signal to Noise ratio

Noise is an important factor impacting the accuracy of the MFMT inverse problem. To describe the noise level of the MFMT system more faithfully, we define the Signal-to-Noise Ratio (SNR) in our experiment as follows:

$$\text{SNR} = \frac{\mu(|S_f - S_b|)}{\sigma(|S_b - S_r|)} \quad (18)$$

where S_f is the fluorescence signal from all source positions, and S_b is the average of S_f with the ROI. S_r is also the fluorescence signal but with the sample replaced by a beam dump. So $\mu(|S_f - S_b|)$ indicate the mean of the true signal intensity, and $\sigma(|S_b - S_r|)$ indicate the standard deviation of the noise due to sample readout.

For *in silico* experiments, a Gaussian noise with mean of zero and standard deviation of SNR is added to the simulated data set to mimic experimental measurements. We use SNR values ranging from 2 to 6 to explore the robustness of the proposed computational approaches on different noise levels.

3.4 Quantitative metrics to assess 3D reconstructions

For consistency, the same normalized figures of merit as [19] are used to quantify the difference between reconstruction results and the ground truth. Four normalized similarity metrics, normalized sum squared difference (nSSD), normalized sum absolute difference (nSAD), normalized disparity (nD), and normalized correlation (nR), are defined as follows.

$$\text{nSSD}(\mathbf{A}, \mathbf{B}) = 1 - \frac{1}{N} \sum_{i=1}^R \sum_{j=1}^C \sum_{k=1}^Z [\mathbf{A}(i, j, k) - \mathbf{B}(i, j, k)]^2 \quad (19)$$

$$\text{nSAD}(\mathbf{A}, \mathbf{B}) = 1 - \frac{1}{N} \sum_{i=1}^R \sum_{j=1}^C \sum_{k=1}^Z |\mathbf{A}(i, j, k) - \mathbf{B}(i, j, k)| \quad (20)$$

$$nR = \frac{\sum_{i=1}^R \sum_{j=1}^C \sum_{k=1}^Z \mathbf{A}(i, j, k) \times \mathbf{B}(i, j, k)}{\sqrt{\sum_{i=1}^R \sum_{j=1}^C \sum_{k=1}^Z \mathbf{A}(i, j, k)^2 \times \sum_{i=1}^R \sum_{j=1}^C \sum_{k=1}^Z \mathbf{B}(i, j, k)^2}} \quad (21)$$

$$nD = 1 - \frac{1}{N} \sum_{i=1}^R \sum_{j=1}^C \sum_{k=1}^Z \mathbf{A}'(i, j, k) \oplus \mathbf{B}'(i, j, k) \quad (22)$$

where $\mathbf{A}(i, j, k)$, $\mathbf{B}(i, j, k)$ are the value of the normalized numerical model and reconstructed result at coordinates (i, j, k) , $N = R \times C \times Z$ is the total voxels in a cubic phantom and R, C, Z are the number of voxels three axes, respectively. \mathbf{A}', \mathbf{B}' are binaries (0 or 1) of \mathbf{A}, \mathbf{B} through thresholding, and \oplus represents the exclusive or (XOR) operation. The values of these four metrics are between 0 (for two absolute different models) and 1 (for two coincident models). In other words, a larger value of the metrics indicates higher similarity between the reconstruction result and the ground truth.

4. *In silico* and experimental validation

4.1 Simulation settings

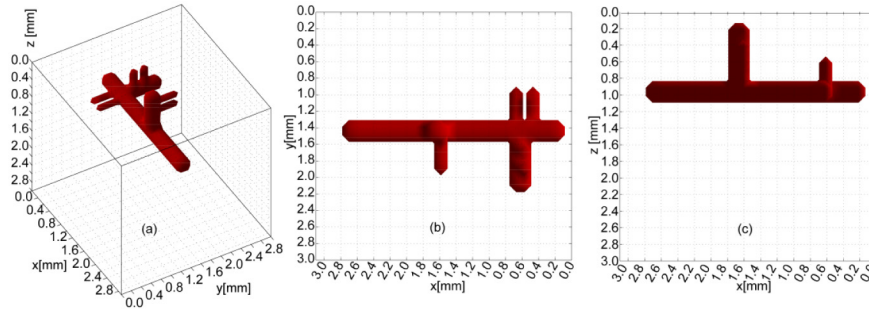


Fig. 2. The numerical phantom designed to mimic a bio-printed vascular channel and sprouting capillaries. (a), (b) and (c) are the full view, xy view, and xz view of the phantom, respectively.

A numerical phantom was designed to evaluate the performance of the proposed postprocessing methods, as shown in Fig. 2. The imaging domain had a surface area of $3.1 \times 3.1 \text{ mm}^2$ with a depth of 3 mm and was uniformly discretized into $31 \times 31 \times 30$ voxels with $100 \text{ }\mu\text{m}$ resolution. A vascular tree with a main trunk and three groups of offshoots was placed within the phantom at $z = 1 \text{ mm}$. The diameter of the trunk and offshoots were $400 \text{ }\mu\text{m}$ and $200 \text{ }\mu\text{m}$ separately, and the separation distance between two adjacent off shoots was $100 \text{ }\mu\text{m}$. The fluorophore concentration was assumed to be uniform in the vessel with effective quantum yield equal to 1. The optical properties of domain were assumed to be homogeneous at the excitation wavelength, with absorption coefficient $\mu_a = 0.02 \text{ mm}^{-1}$, scattering coefficient $\mu_s' = 1 \text{ mm}^{-1}$, index of refraction $n = 1.34$, and anisotropy factor $g = 0.81$. These values are derived from the collagen scaffold typically employed in our bio-printing application at 6-9 mg/ml density [32] and μ_s' is also on the same level with many biological tissues [33].

We replicated the imaging system configuration as used in a real experiment: 961 scanning positions and 48 detectors at each source location. The *in silico* measurements are generated by multiplying the sensitivity matrix with the bio-printed vasculature model. We added Gaussian noise with SNR as low as 2. Then reconstructions are performed as described in Section 2. For each of the following simulations and experiment cases, the optimal regularization parameter is chosen through L-curve analysis [34].

4.2 Coherence reduction via preconditioning

We first evaluated the performance of the preconditioning method on improving the incoherence of the sensitivity matrix and reconstruction results with the numerical phantom. As we mentioned in Section 2.3, it could be directly applied on the whole sensitivity matrix (whole-pre) or block by block (sub-pre). We compared the distribution of largest 30% of normalized inner products between two columns of the sensitivity matrix in Fig. 3(a), and the cumulative coherence of the sensitivity matrix as a function of column number k in Fig. 3(b). As shown in the insert in Fig. 3(a), the sub-pre reduces the relative *Area Under the Curve* (AUC) to 45.27% while whole-pre reduces the AUC to 34.79%. The average slope listed in the insert in Fig. 3(b), a quantitative index of cumulative coherence, also shows a slower increment after applying the preconditioner on whole sensitivity matrix and sub sensitivity matrix, which demonstrates both preconditioning could effectively improve the orthogonality of the matrix.

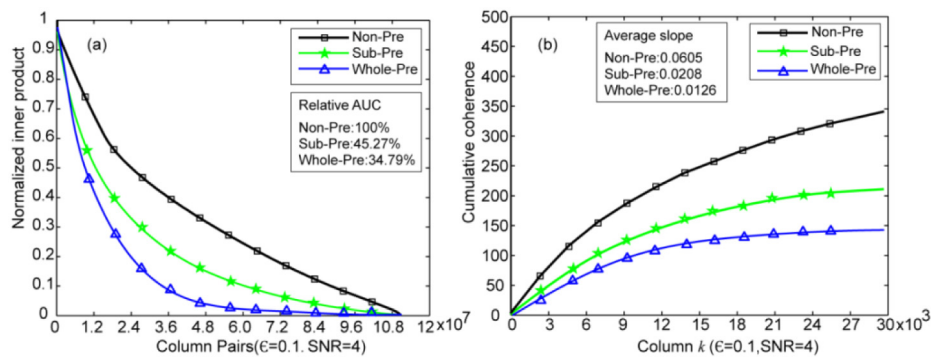


Fig. 3. Curves of the normalized products (a) and cumulative coherence (b) before and after applying the preconditioner on whole sensitivity matrix and block by block sensitivity matrix.

We then compared the reconstruction results between non-pre, whole-pre and sub-pre, at different measurement noise levels (from 2 to 6). The comparison of four reconstruction metrics between non-pre and whole-pre is shown in Fig. 4(a), and those between whole-pre and sub-pre is shown in Fig. 4(b). At all noise levels, the reconstruction quality with the proposed preconditioning methods outperforms that without preconditioning. Moreover, though less effective in reducing the coherence of sensitivity matrix, sub-pre can retrieve the *in silico* model with greater similarity and smaller disparity compared to whole-pre. The 3D reconstruction images of three cases are shown in Fig. 4(c)–4(e). In accordance with the evaluation metrics, sub-preconditioning provides visually the best result. Although whole preconditioning is better at reducing matrix coherence, the reconstruction results are less ideal than sub preconditioning. This observation is in agreement with the results in [18] and [16], which can be explained by the higher condition number and thus more severe noise amplification when applying whole-preconditioning.

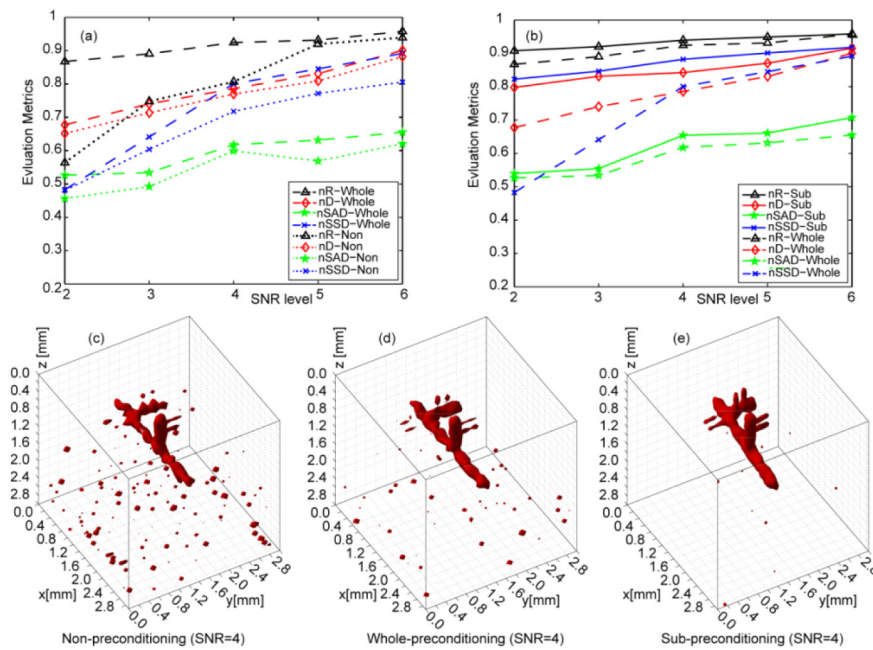


Fig. 4. Reconstruction results and evaluation metrics before and after applying preconditioning. (a) shows the metrics of reconstructions of non-pre and whole-pre at different SNR levels. (b) shows the metrics of reconstructions of whole-pre and sub-pre at different SNR levels. (c-e) display the 3-D reconstruction of non-pre, whole-pre and sub-pre, respectively.

4.3 Faster convergence via the weighted iteration improvement method

We further evaluated the performance of the proposed Weighted Iteration Improvement Method (WIIM) on accelerating the convergence rate of the inverse problem in four scenarios, named whole-pre without WIIM (M1), sub-pre without WIIM (M2), whole-pre with WIIM (M3) and sub-pre with WIIM (M4). It can be seen from Figs. 5(a)–5(c) that the convergence rates without WIIM always stagnate before the desired tolerance, especially for the cases with low SNR level. However, after our l_1 -norm regularization based iterative algorithm is embedded with WIIM, an effective improvement on the convergence rate can be seen. The last residual norm $\|r^k\|$ always becomes much smaller, as reported in Table 1. By comparing M1 with M3 (both whole-preconditioning) and M2 with M4 (both sub-preconditioning), we can see from Table 1 that WIIM could help improve reconstruction results for both Jacobians. On the other hand, the evaluation metrics of M3 worse than M2 shows that preconditioning has greater impact on the results than WIIM because the latter one is just a minor improvement for the iterative optimization algorithm. In addition, sub-pre converges faster than whole-pre when WIIM is applied, as illustrated by the green and blue curves in Fig. 5(a)–(c). Overall, the proposed WIIM works well for both whole sensitivity matrix and sub sensitivity matrix to fasten convergence rate.

The visual reconstruction results in Fig. 5(d)–(f) further demonstrate the improvement of reconstruction quality with WIIM. The obvious improvement from Fig. 5(d)–5(f) indicates that the deterioration of reconstruction could likely results from the stagnation before the iterative algorithm converges to the desired tolerance.

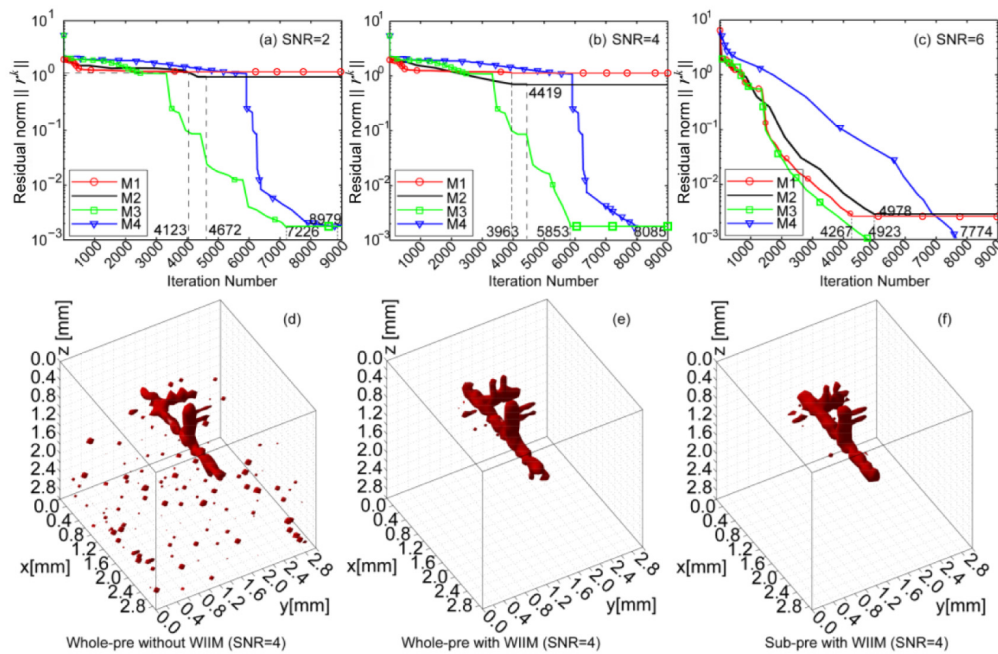


Fig. 5. Comparison of convergence rate and visual reconstruction results among four scenarios without and with WIIM embedded in them. (a-c) plots the convergence rate and iteration number of the scenarios at three noise levels, respectively. (d-f) visualize the reconstruction of three scenarios at SNR of 4, respectively.

Table 1. Quantification results of the preconditioning for numerical data (Targeted tolerance = 10^{-3} , Preset max Iteration = 9,000)

SNR	Method	Condition number	Residual norm	Number of Iteration	nR	nD	nSAD	nSSD
2	whole-pre without WIIM (M1)	256.42	1.2207	4,123	0.8654	0.6798	0.5245	0.4880
	sub-pre without WIIM (M2)	178.29	0.9656	4,672	0.9027	0.8091	0.5488	0.8223
	whole-pre with WIIM (M3)	42.76	0.1017	8,979	0.8727	0.6910	0.5476	0.4920
	sub-pre with WIIM (M4)	25.19	0.1017	7,226	0.9129	0.8221	0.5686	0.8316
4	whole-pre without WIIM (M1)	256.42	1.1479	3,963	0.9143	0.7822	0.6134	0.7908
	sub-pre without WIIM (M2)	178.29	0.8654	4,419	0.9327	0.8398	0.6577	0.8674
	whole-pre with WIIM (M3)	42.76	0.0081	8,085	0.9278	0.8091	0.6930	0.8242
	sub-pre with WIIM (M4)	25.19	0.0081	5,853	0.9429	0.8510	0.7123	0.8821
6	whole-pre without WIIM (M1)	256.42	0.0028	4,267	0.9443	0.8901	0.6576	0.8951
	sub-pre without WIIM (M2)	178.29	0.0026	4,978	0.9443	0.9218	0.7005	0.9241
	whole-pre with WIIM (M3)	42.76	0.00089	7,774	0.9528	0.9091	0.7212	0.9258
	sub-pre with WIIM (M4)	25.19	0.00089	4,923	0.9619	0.9315	0.7406	0.9367

4.4 Performance of preconditioning methods on experimental data

Next we evaluated the performance of the proposed preconditioning and regularization methods on an experimental data set. The collagen phantom to be explored is homogeneous with optical properties $\mu_a = 0.02 \text{ mm}^{-1}$, $\mu_s' = 1 \text{ mm}^{-1}$, $n = 1.34$, and $g = 0.81$. It has a size of $3.1 \times 3.1 \times 3.0 \text{ mm}^3$, and voxels of $100 \times 100 \times 100 \text{ }\mu\text{m}^3$, same as the simulated case. Four polystyrene fluorophore beads (GFP 488/509, Cospheric) were placed 1.7~1.8 mm beneath

the sample surface under standard protocols. The transversal slices across the x-y and y-z planes taken from a micro-MRI are shown in Figs. 6(a)–(b) and the whole 3-D view of the phantom, overlaid with the best reconstruction result, is shown in Fig. 6(c).

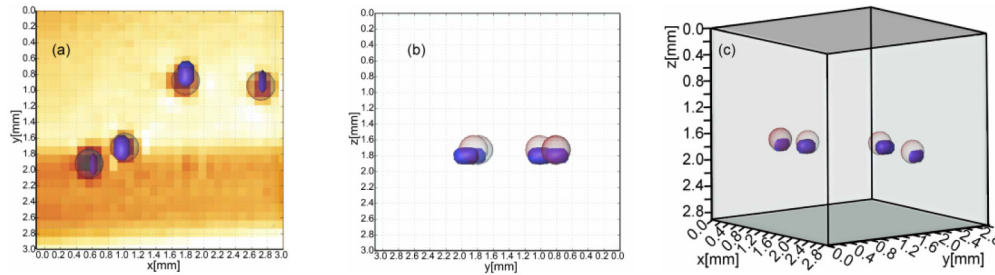


Fig. 6. Experimental phantom (a-c) are the x-y, x-z cross-sectional, and 3D view of phantom obtained from micro-MRI, overlaid with the best reconstruction result, respectively.

The measurements were recorded as described in Section 2.1 and the data acquisition time at each scanning position was less than 20 ms. Similar to the *in silico* cases, we compared the condition number, convergence rate, and reconstruction metrics, when no / whole / sub-preconditioning is preformed, with and without WIIM. All the inverse problems were solved with our l_1 -norm reconstruction algorithm as described in Section 2.2 and the optimal regularization parameters were determined through L-curves. As an example, Fig. 7(a) shows the L-curve for experimental data to determine the optimal regularization parameter when applying sub-preconditioning.

Table 2. Quantification results of the preconditioning for experimental data (Targeted tolerance = 10^{-3} , Preset max Iteration = 9,000, SNR = 4)

Method	Condition number	Relative AUC	Residual norm	Number of Iteration	nR	nD	nSAD	nSSD
non-preconditioning	648,496	100%	1.1479	3,963	0.5440	0.6517	0.4562	0.4809
whole-pre without WIIM	268.91	34.79%	1.1426	7,910	0.9079	0.7762	0.6106	0.7902
sub-pre without WIIM	189.75	45.27%	1.1437	5,792	0.9270	0.8280	0.6513	0.8593
whole-pre with WIIM	49.08	34.79%	0.0081	8,085	0.9173	0.8075	0.6898	0.8217
sub-pre with WIIM	26.85	45.27%	0.0081	5,853	0.9379	0.8478	0.7065	0.8712

As shown in Table 2, the proposed preconditioner can effectively reduce the condition number as well as the coherence between columns of the sensitivity matrix in both formations of. It leads to a reduction of 2,524 and 3,628 times of condition number in the two sensitivity matrix formations, respectively. Meanwhile, the last residual norm also drops from 1.1479 to 0.0081 with the help of WIIM, which is very close to the targeted tolerance. The four evaluation metrics of reconstruction results in Table 2 further indicate the proposed computational methods greatly improve the reconstruction fidelity. Although whole-pre and sub-pre strategies have the same last residual norm, sub-pre achieves lower condition number as well as faster convergence rate. Visual results in Fig. 7(b)–7(e) and evaluation metrics in Table 2 also validate the superiority of reconstruction quality from sub-pre. However, as shown in Figs. 7(d), even the best results still suffer from some artifact, which illustrates that the proposed methods may still be sensitive to noise. Figure 7(e) shows the reconstruction result when applying both methods on the sub sensitivity matrices and noise suppression strategy as described in [19], which is very close to the ground truth obtained from micro-MRI.

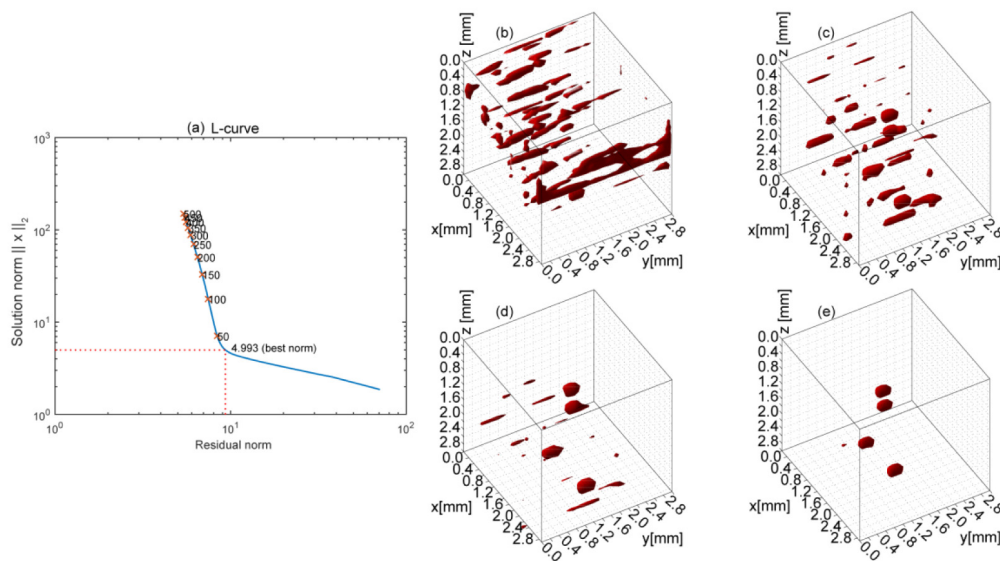


Fig. 7. Reconstruction results of experimental phantom. (a) L-curve for experimental data to determine the optimal regularization parameter when applying sub-preconditioning. (b-d) are results without preconditioning, with whole-pre and WIIM, and with sub-pre and WIIM, respectively. (e) is the result when applying sub-pre, WIIM and noise suppression on the experimental phantom.

5. Conclusion

We have proposed two computational methods to reduce the coherence and condition number of the sensitivity matrix of MFMT, and to accelerate convergence rate of the adopted iterative algorithm when solving the inverse problem. The preconditioning method can be directly applied on the whole sensitivity matrix or on sub-matrices of each detector, while the sub-preconditioning strategy provides fast convergence rate and better reconstruction quality than whole-preconditioning despite larger coherence between columns of the sensitivity matrix. This is in agreement with the observations in [16–18], because direct preconditioning usually lead to a large condition number and amplification of noise in the measurement vector. The regularization method based on WIIM proves to be effective to avoid stagnation before the iterative algorithm reaches the target tolerance. We have tested the performance of these two methods on *in silico* as well as *in vitro* data sets. In both cases, the reconstruction fidelity is significantly improved compared to previous results. We plan to further investigate the potential of the proposed methods for augmented data sets, such as in time-resolved cases [26] when applied to FRET tomography [35] to enable to monitor cellular processes [36,37] and in the case of phased array systems [38–41].

Funding

National Institutes of Health (NIH) (R01-EB019443 and R01 BRG-CA20772); Natural Science Foundation of Shandong Province (ZR2018MF034); National Natural Science Foundation of China (61472227 and 81771532).

Disclosures

The authors declare that there are no conflicts of interest related to this article.